Multihistogram Reweighting for Nonequilibrium Markov Processes Using Sequential Importance Sampling Methods

Troels Arnfred Bojesen*

Department of Physics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

(Dated: April 8, 2013)

We present a multihistogram reweighting technique for nonequilibrium Markov Chains with discrete energies. The method generalizes the single histogram method of Yin et al. [Phys. Rev. E **72**, 036122 (2005)], making it possible to calculate the time evolution of observables at a posteriori chosen couplings based on a set of simulations performed at other couplings. In the same way as multihistogram reweighting in an equilibrium setting improves the practical reweighting range as well as use of available data compared to single histogram reweighting, the method generalizes the multihistogram advantages to nonequilibrium simulations. We demonstrate the procedure for the Ising model with Metropolis dynamics, but stress that the method is generally applicable to a range of models and Monte Carlo update schemes.

I. INTRODUCTION

In the last two decades the use of short time critical dynamics (STCD) as a way of investigating the critical properties of models in statistical mechanics has emerged as an interesting alternative to equilibrium Monte Carlo (EMC) simulations. The reason is twofold: First, STCD simulations may, given sufficiently large system sizes, effectively avoid finite size effects, since the correlation length is still small in the short time regime. Second, and even more important, critical slowing down is avoided due to the system being far from equilibrium. See for instance Refs. 1 and 2 for some recent reviews on the STCD method.

The EMC simulations have, on the other hand, long had the advantage of the powerful histogram reweighting methods of Ferrenberg and Swendsen, meaning that knowledge of the exact couplings of interest are not needed a priori the simulations: Single histogram reweighting [3] makes it, in principle, possible to extract observables at arbitrary couplings from a simulation performed at a fixed coupling, although the practical range is limited to a small neighborhood of the simulation coupling. The practical reweighting range as well as the statistical quality of the observables can be improved by using multihistogram reweighting [4], where data from an arbitrary number of simulations at different couplings are combined in an optimal way.

A first attempt to introduce reweighting techniques to STCD simulations was done by Lee and Okabe [5]. They proposed what essentially is (Monte Carlo) time dependent single histogram reweighting, where the reweighting is done simultaneously with the simulation. This is a major drawback of the method, as all the desired couplings, simulation as well as reweighted, must be known a priori. In addition, obtaining an observable through reweighting is almost as computationally heavy as performing the simulation itself, and since the procedure scales linearly with the number of reweighting couplings, the practical usefulness of the method is limited.

Yin et al. proposed a novel method that improves the single histogram reweighting for STCD, but the method is restricted to models with a discrete energy spectrum [6]: By keeping track of a time dependent histogram of the energy changes at each step in the Markov chain, and sampling these each time an observable is sampled, it is possible to efficiently reweight observables to a posteriori chosen couplings. This is a significant improvement, but the reweighting is still limited to the single histogram regime where only data from simulations at one coupling strength is used. In a real world situation one would typically perform several simulations at different couplings when searching for the critical one, thus obtaining a lot of data that contains potentially useful information. It would therefore be advantageous to develop a multihistogram reweighting technique for nonequilibrium Markov processes.

In this paper we propose such a technique for nonequilibrium reweighting of models with *discrete energy spectra* by extending the method of Yin et al. to time dependent multihistogram reweighting using the method of Ferrenberg and Swendsen. After the general derivation we test the method on a Metropolis dynamics STCD simulation on the 2D Ising model and compare it with the "raw" (non-reweighted) as well as single histogram reweighted results.

Note that although these methods are developed with STCD simulations in mind, they are generally applicable to all nonequilibrium Markov processes (and hence also equilibrium processes).

II. MULTIHISTOGRAM REWEIGHTING

A. General Formulation

Let $x_t = (\sigma_1, \sigma_2, \ldots, \sigma_t)$ denote a Markov chain of field configurations σ after t steps (often called "time" from now on), and let Σ_t be the dynamical phase space, i.e. the

* troels.bojesen@ntnu.no

set of all *possible* Markov chains of t steps, starting from an initial set of states Σ_0. Given a (usually small) set of coupling parameters β, we may associate a probability weight to each Markov chain:

$$w : \beta \times \Sigma_t \to \mathbb{R}. \qquad (1)$$

At each time step we have a probability distribution over all possible *reachable* states, which we may capture in the *dynamical* partition function

$$Z_t(\beta) \equiv \sum_{x_t} w(\beta, x_t), \qquad (2)$$

where x_t under the summation sign is a shorthand notation for $x_t \in \Sigma_t$. In the same way we define the dynamical average of an observable O as

$$\langle O \rangle_t (\beta) \equiv Z_t(\beta)^{-1} \sum_{x_t} w(\beta, x_t) O(\sigma_t). \qquad (3)$$

Thus, at a fixed time t there is no difference between the dynamical formalism and the equilibrium formalism, except that the fundamental "states" in the dynamical case are entire Markov chains instead of single field configurations. With this in mind, methods from equilibrium simulations may be carried right over to dynamic simulations.

For the sake of simplicity, we will from now on restrict ourselves to the special case when the set of coupling parameters consists of only one coupling parameter (also named β). Generalization to multiple parameters is straightforward.

In the equilibrium formalism, Ferrenberg–Swendsen multihistogram reweighting goes as follows [4]: Let $H(\sigma) = H_0(\sigma) + \beta S(\sigma)$ be a general Hamiltonian of a field σ and the partition function $Z(\beta) = \sum_\sigma \exp(H(\sigma))$. S is an operator (energy, magnetization, etc.) on the field. Given Q EMC simulations performed at couplings $\{\beta_1, \ldots, \beta_q, \ldots, \beta_Q\}$, each with N_q samples with autocorrelation time τ_q, the best approximation for the probability weight of S at a (a posteriori chosen) coupling β^* is given by

$$P(\beta^*, S) = \frac{\sum_q (1 + 2\tau_q)^{-1} M_q(S) \exp(\beta^* S)}{\sum_q N_q (1 + 2\tau_q)^{-1} \exp(\beta_q S) \tilde{Z}(\beta_q)^{-1}}, \qquad (4)$$

where $M_q(S)$ is the histogram of the S-measurements of the q'th series, and

$$\tilde{Z}(\beta) = \sum_S P(\beta, S). \qquad (5)$$

P is found by solving Eqs. (4) and (5) iteratively. The error minimized average of an observable (operator on S) at β^* is then given by

$$\langle O \rangle (\beta^*) = \tilde{Z}(\beta^*)^{-1} \sum_S O(S) P(\beta^*, S). \qquad (6)$$

We now turn to the dynamic generalization of this method. Let $X_{q,t} \equiv \{x_{q,t}\}$ be a set of N_q Markov chains obtained from independent simulations at the same coupling β_q, and let $X_t \equiv \{X_{q,t} \mid q \in \{1, \ldots, Q\}\}$ be the collection of all Markov chains from Q such sets, each with their own associated coupling $\beta_q \in \{\beta_1, \ldots, \beta_Q\}$. Replacing $\exp(\beta S) \to w(\beta, x_t)$, $S \to x_t$ (i.e. x_t may be seen as the identity operator), $P(\beta, S) \to W(\beta, x_t)$, and $\tilde{Z}(\beta) \to \tilde{Z}_t(\beta) = \sum_{x_t \in X_t} W(\beta, x_t)$ in Eq. (4), we get

$$W(\beta^*, x_t) = \frac{\sum_q M_q(x_t) w(\beta^*, x_t)}{\sum_q N_q w(\beta_q, x_t) \tilde{Z}_t(\beta_q)^{-1}}. \qquad (7)$$

$\tau_q = 0$ since the Markov chains are independent. $M_q(x_t)$ is the histogram of Markov chains on the form x_t for $X_{q,t}$, i.e. $M_q(x_t) = \sum_{y \in X_{q,t}} \delta_{y,x_t}$. Now, if we treat all the obtained Markov chains as unique (even in the improbable case of two being identical), there will be *only one* $y \in X_t$ fulfilling $y = x_t$, namely x_t itself. Hence we must have that $\sum_q M_q(x_t) = 1$, and Eq. (7) simplifies to

$$W(\beta^*, x_t) = \frac{w(\beta^*, x_t)}{\sum_q N_q w(\beta_q, x_t) \tilde{Z}_t(\beta_q)^{-1}}. \qquad (8)$$

The error minimizing dynamical average is then

$$\langle O \rangle_t (\beta^*) = \tilde{Z}_t(\beta^*)^{-1} \sum_{x_t} O(\sigma_t) W(\beta^*, x_t), \qquad (9)$$

which is what we are interested in. To get any further with a practical use of Eqs. (8) and (9) we need to determine the Markov chain weights $\{w(\beta, x_t)\}$.

B. The Markov Chain Weight

The weight of a Markov chain [Eq. (1)] is per definition given by

$$w(\beta, x_t) = \omega(\sigma_0) \prod_t \omega(\beta, \sigma_t | \sigma_{t-1}), \qquad (10)$$

where $\omega(\sigma_0)$ is the weight of the initial state and $\omega(\beta, \sigma_t | \sigma_{t-1})$ is the weight associated with a Markov step from σ_{t-1} to σ_t at coupling β. Since each Markov chain in an STCD simulation may consist of a very large number of steps (e.g. 10^{10} for a 10^4 sweep simulation on a 1000×1000 lattice), it is practically impossible to directly store the weight information of all steps for later reweighting; in the general case, where the set of possible step weights is large (possibly infinite), one has no choice but to reweight "on the fly", as Lee and Okabe proposed.

If, however, we restrict ourselves to models where the set of possible step weights is discrete, it *is* possible to find the the Markov chain weight at an a posteriori chosen coupling β. The trick is to write Eq. (10) as [6]

$$w(\beta, x_t) = \omega(\sigma_0) \prod_\alpha \omega(\beta, \alpha)^{r_{\alpha,t}}, \qquad (11)$$

$\omega(\beta, \alpha)$ being the weight of a step of type α, and $r_{\alpha,t} \in \mathbb{N}_0$ the number of such steps in the Markov chain x_t. $\omega(\beta, \alpha)$ can be calculated for an arbitrary β, given the knowledge of α and the dynamics used in the simulation. Thus, to know $w(\beta, x_t)$ we only need to know the initial state σ_0 and the histogram of $r_{\alpha,t}$-values.

As an example, consider an STCD simulation of a spin model using canonical Metropolis dynamics at some coupling β_q. A trial update of a spin will result in an energy change ΔE_α, which we assume to be member of a discrete set of all possible (one step) energy changes $\{\Delta E\}$. There are three possible outcomes of a trial update in the Metropolis dynamics: (See Ref. 5 for mathematical details.)

- $\Delta E_\alpha \leq 0$, in which case the trial is always accepted. Then $\omega(\beta, \alpha) = 1$.

- $\Delta E_\alpha > 0$ and the update is accepted. Then $\omega(\beta, \alpha) = \exp(-\beta \Delta E_\alpha)$.

- $\Delta E_\alpha > 0$ and the update is rejected. Then $\omega(\beta, \alpha) = 1 - \exp(-\beta \Delta E_\alpha)$.

The explicit expression for the Markov chain weight, Eq. (11), becomes

$$
\begin{aligned}
w(\beta, x_t) &= \omega(\sigma_0) \prod_{\substack{\alpha \\ \Delta E_\alpha > 0 \\ \text{accept}}} [\exp(-\beta \Delta E_\alpha)]^{r_{\alpha,t}} \\
&\times \prod_{\substack{\alpha \\ \Delta E_\alpha > 0 \\ \text{reject}}} [1 - \exp(-\beta \Delta E_\alpha)]^{r_{\alpha,t}} \\
&= \omega(\sigma_0) \exp(-\beta \Delta_t E^{\mathrm{a}}) \\
&\times \prod_{\substack{\alpha \\ \Delta E_\alpha > 0 \\ \text{reject}}} [1 - \exp(-\beta \Delta E_\alpha)]^{r_{\alpha,t}},
\end{aligned} \quad (12)
$$

where

$$
\Delta_t E^{\mathrm{a}} \equiv \sum_{\substack{\alpha \\ \Delta E_\alpha > 0 \\ \text{accept}}} r_{\alpha,t} \Delta E_\alpha = \sum_{\substack{t \\ \Delta_t E > 0 \\ \text{accept}}} \Delta_t E \quad (13)
$$

is the running sum of all positive acceptance energy changes. So, to be able to calculate the Markov chain Metropolis dynamics weight at an arbitrary β, and thus to be able to multihistogram reweight an observable average using Eq. (9), one has to keep track of $\Delta_t E^{\mathrm{a}}$ as well as a histogram $r_{\alpha,t}$ over the distribution of the rejected positive energy changes. For many discrete models this is a manageably small set of extra information to sample in addition to the standard observables.

III. TESTING THE METHOD ON THE $2D$ ISING MODEL

We test the nonequilibrium multihistogram reweighting method on a Metropolis dynamics STCD simulation
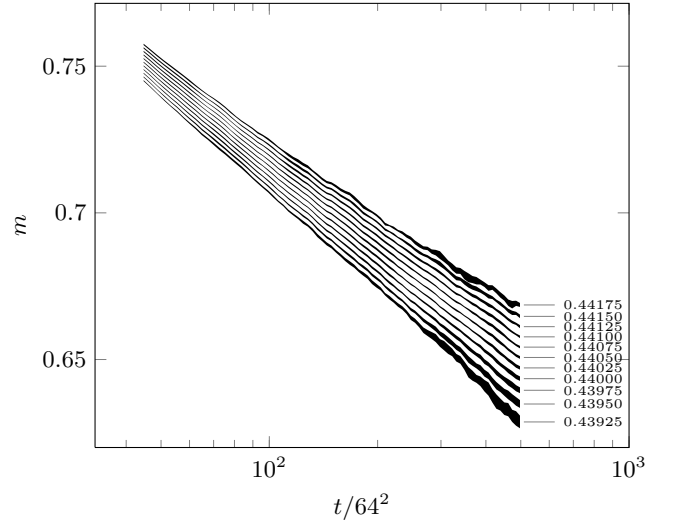

FIG. 1. Plot of magnetization curves for the multihistogram reweighted couplings (from below) $\beta^* = 0.439\,25, 0.439\,50, \ldots, 0.441\,75$. The reweighting is based on simulations at $\beta_q = 0.4400, 0.4405$ and 0.4410. Line thickness corresponds to $\pm$ jackknife standard error.

of a two dimensional ferrormagnetic Ising model on a square $L \times L = 64 \times 64$ lattice with periodic boundary conditions. The Hamiltonian is given by

$$
H[s] = -\sum_{\langle i,j \rangle} s_i s_j, \quad (14)
$$

where i and j are lattice indices and $s_i \in \{-1, 1\}$. Note that $\{\Delta E\} = \{-4, -2, 0, 2, 4\}$, which means that we just need to sample three extra quantities in addition to the observable(s), in this case the dynamical average of the magnetization per site,

$$
m_t \equiv L^{-2} \sum_i s_{t,i}. \quad (15)
$$

To make the test simple, we consider only simulations from a perfectly ordered state, $\Sigma_0 = \sigma_0 = \{s_i = 1 \mid \forall i\}$, and so we may choose $\omega(\sigma_0) = 1$.

Figure 1 shows a fan of reweighted magnetization curves as a function of Monte Carlo time, obtained from three simulations performed at $\beta_1 = 0.4400$, $\beta_2 = 0.4405$, and $\beta_3 = 0.4410$. (This is close to criticality, $\beta_c = \ln(1 + \sqrt{2})/2 = 0.440\,068\,7\ldots$) $N_q = 20\,000\ \forall q$. Errors are obtained by the Jackknife method. Notice how the error grows as the reweighted coupling deviates more from the simulated couplings and the overlap in energy histograms worsens, as is the case for all statistical reweighting techniques.

We compare the multihistogram reweighting with raw data and single histogram reweighting. First we perform single histogram reweighting of the $q = 1$ simulation to $\beta^* = \beta_2$ and from simulation 3 to $\beta^* = \beta_2$, then multihistogram reweighting to $\beta^* = \beta_2$ using simulation 1 and 2,
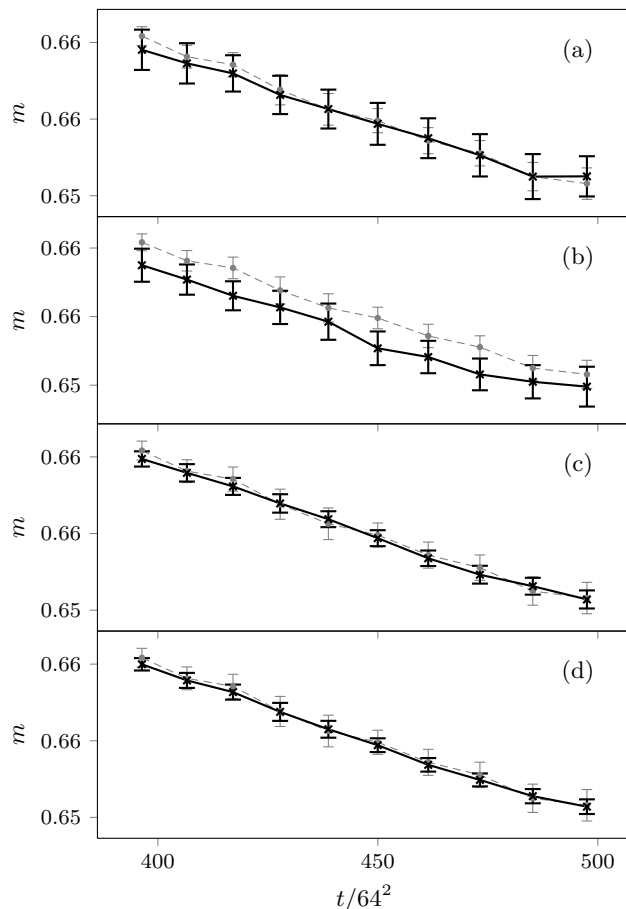
FIG. 2. Comparison between non-reweighted magnetizations curves (dashed gray lines) at $\beta = 0.4405$ and reweighted curves (thick black lines), reweighted to $\beta^* = 0.4405$. (a) shows the single histogram reweighting from a simulation at $\beta_q = 0.4400$. (b) shows the single histogram reweighting from a simulation at $\beta_q = 0.4410$. (c) shows a multihistogram reweighting based on both these simulations. (d) shows a multihistogram reweighting using all data available, i.e. the simulations as $\beta_q = 0.4400$, 0.4405 and 0.4410. Error bars are jackknife standard errors.

and finally multihistogram reweighting to $\beta^* = \beta_2$ using the entire set of available data, $q = 1, 2, 3$. The results are compared with the non-reweighted dynamical average of the raw data of simulation 2 in Fig. 2. Notice that while the single histogram reweighting always performs worse (in terms larger error bars) than the raw data, the multihistogram reweighting may even outperform the raw data, given that, like in this case, more simulations with overlapping energy histograms exist.

It should be noted that the computational cost of multihistogram reweighting is of the same order of magnitude as single histogram reweighting – and negligible compared to the simulations.

IV. CONCLUSION

To summarize, we have generalized the single histogram nonequilibrium Markov Chain reweighting technique of Yin et al. [6] to a multihistogram framework based on the equilibrium method of Ferrenberg and Swendsen. In this manner we can take advantage of several nonequilibirum simulations performed at different (but close) couplings, producing dynamical averages with equal or smaller errors than those obtained by averages based on the non-reweighted or single histogram reweighted datasets.

[1] Y. Ozeki and N. Ito, Journal of Physics A: Mathematical and Theoretical **40**, R149 (2007).

[2] E. V. Albano, M. A. Bab, G. Baglietto, R. A. Borzi, T. S. Grigera, E. S. Loscar, D. E. Rodriguez, M. L. R. Puzzo, and G. P. Saracco, Reports on Progress in Physics **74**, 026501 (2011).

[3] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).

[4] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).

[5] H. K. Lee and Y. Okabe, Phys. Rev. E **71**, 015102 (2005).

[6] J. Q. Yin, B. Zheng, and S. Trimper, Phys. Rev. E **72**, 036122 (2005).